Integration of transcriptome and genome sequencing uncovers functional variation in human populations

Functional interpretation of genetic variants discovered in human genomes sequencing is essential to understand human phenotypic variation. In a study of the Geuvadis project, we sequenced mRNA and small RNA of LCLs from 465 individuals from CEU, TSI, GBR, FIN and YRI populations of the 1000 Genomes samples. Distributing the sequencing across seven laboratories with a large set of replicates first allows us to outline best practices and assess technical variability of large-scale transcriptome sequencing. We also develop methods for variant calling from RNA-seq data. Furthermore, we analyze the variation and interaction of transcript and miRNA quantities and alternative splicing between individuals and populations in an unprecedented scale. The combination of transcriptome and genome sequencing data gives us the opportunity to characterize both common and rare regulatory variants and their interactions. These data together with the rich functional annotation of the genome shed light on the functional mechanisms of regulatory variation and bring us closer to causal regulatory variants. This study takes us toward not only cataloguing regulatory variants but also understanding and even predicting how variants in the human genome affect multiple layers of the transcriptome.

Tuuli Lappalainen 1

Michael Sammeth 2,3

Natalja Kurbatova 4

Tim Strom 5

Peter ‘t Hoen 6

Matthias Barann 7

Marc Sultan 8

Esther Lizano 2

Manuel Rivas 9

Alvis Brazma 4

Ann-Christine Syvanen 10

Thomas Meitinger 5

Philip Rosenstiel 7

Ralf Sudbrak 8

Ivo Gut 3

Xavier Estivill 2

Emmanouil Dermitzakis 1

tuuli.lappalainen@unige.ch

micha@sammeth.net

natalja@ebi.ac.uk

TimStrom@helmholtz-muenchen.de

P.A.C.\_t\_Hoen@lumc.nl

m.barann@ikmb.uni-kiel.de

sultan@molgen.mpg.de

esther.lizano@crg.es

rivas@broadinstitute.org

brazma@ebi.ac.uk

ann-christine.syvanen@medsci.uu.se

meitinger@helmholtz-muenchen.de

p.rosenstiel@mucosa.de

sudbrak@molgen.mpg.de

igut@pcb.ub.es

xavier.estivill@crg.eu

dermitza@unige.ch

1

University of Geneva

Dept of Genetic Medicine and Development

rue Michel Servet 1

1211

Geneva

2

Ctr for Genomic Regulation and UPF

Dr. Aiguader 88

08003

Barcelona

3

Centro Nacional de Analisis Genomico

Baldiri Reixac 4

08028

Barcelona

4

European Bioinformatics Institute

Wellcome Trust Genome Campus

Hinxton

CB10 1SD

5

Helmholtz Zentrum München

Inst of Human Genetics

Ingolstädter Landstr. 1

85764

Neuherberg

Germany

6

Leiden University Medical Ctr

Ctr for Human and Clinical Genetics

Postzone S4-P PO Box 9600

2300 RC

Leiden

7

University of Kiel

Institute of Clinical Molecular Biology

Schittenhelmstr.12

24105

Kiel

8

Max Planck Institute for Molecular Genetics

Dept of Vertebrate Genomics

Ihnestr. 63 - 73

14195

Berlin

9

Wellcome Trust Centre for Human Genetics

Roosevelt Drive

Oxford

OX3 7BN

10

Uppsala University

Dept of Medical Sciences

Research dept. II University Hospital

75185

Uppsala